AFRL-SA-WP-SR-2013-0005

Factor Analysis of Computer-Based Multidimensional Aptitude Battery-Second Edition Intelligence Testing from Rated U.S. Air Force Pilots

Wayne L. Chappelle
Bret D. Heerema
U.S. Air Force School of Aerospace Medicine, Wright-Patterson AFB, OH

William T. Thompson
NeuroStat Analytical Solutions, LLC, San Antonio, TX

September 2012

Final Report
for July 2010 to June 2011

Distribution A: Approved for public release; distribution is unlimited.
Case Number: 88ABW-2013-1912,
22 Apr 2013

Air Force Research Laboratory
711th Human Performance Wing
School of Aerospace Medicine
Aerospace Medicine Department
Neuropsychiatry Branch
2510 Fifth St.
Wright-Patterson AFB, OH 45433-7913

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

Qualified requestors may obtain copies of this report from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-SA-WP-SR-2013-0005 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNATURE//	//SIGNATURE//
COL LEE BEYER Chief, Aerospace Med Consult Div	COL ROBERT E. CARROLL Chair, Aerospace Medicine Dept

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*
1 Sep 2012

2. REPORT TYPE
Special Report

3. DATES COVERED *(From – To)*
July 2010 – June 2011

4. TITLE AND SUBTITLE

Factor Analysis of Computer-Based Multidimensional Aptitude Battery-Second Edition Intelligence Testing from Rated U.S. Air Force Pilots

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Wayne L. Chappelle, Bret D. Heerema, William T. Thompson

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
USAF School of Aerospace Medicine
Aerospace Medicine Department/FECN
2510 Fifth St.
Wright-Patterson AFB, OH 45433-7913

8. PERFORMING ORGANIZATION REPORT NUMBER

AFRL-SA-WP-SR-2013-0005

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The purpose of this study is to assess the factorial structure regarding a commercially available computer-based intelligence test given to U.S. Air Force (USAF) pilot training candidates (manned as well as unmanned) during medical flight screening. Principal components analysis was conducted on the intelligence test scores from the computer-based Multidimensional Aptitude Battery-Second Edition (MAB-II). This test was administered to 10,612 USAF pilot training candidates. Subtest and measurement model correlations were estimated. Confirmatory factor analysis using the MAB-II showed positive correlations between specific subtests. The results of factor analyses suggest MAB-II intelligence test scores are best suited to a three-factor model unique to the rated USAF pilot population. In addition to verbal- and performance-based intelligence quotients, a visual reasoning and working memory intelligence quotient composed of the arithmetic, digit symbol, and spatial analyses subtests appears in test scores for this population. This finding is in contrast to the two-factor measurement model used for the general population. The results of the study support an alternative way for conceptualizing and reorganizing MAB-II intelligence test subtest scores to better account for the relationship between test scores from USAF pilot training candidates. The relationship of these test scores must be well understood to effectively evaluate how specific cognitive aptitudes are related to (and/or affected by) changes in any particular subtest. Rather than just measure global cognitive ability, it is recommended that neuropsychologists and clinical psychologists remain sensitive to the pattern and factor analyses of subtests that comprise the global score of intelligence for this commercially available intelligence test. This is especially important when the purpose of a psychological evaluation is to assess the cognitive strengths and weaknesses of a specific individual (and his or her readiness to participate) in the unique high-risk, high-demand nature of USAF military pilot duties.

15. SUBJECT TERMS
Neuropsychological screening, factorial structure, intelligence testing, pilots, medical flight screening

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Wayne L. Chappelle
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	SAR	19	

This page intentionally left blank.

TABLE OF CONTENTS

Section		Page
1.0	SUMMARY	1
2.0	INTRODUCTION	1
	2.1 Conceptualizing and Measuring Intelligence	2
	2.2 Purpose of the Study	3
3.0	METHODS	4
	3.1 Participants	4
	3.2 Instrument	4
	3.3 Procedure	4
4.0	RESULTS	6
5.0	DISCUSSION	9
6.0	CONCLUSIONS	10
7.0	REFERENCES	11
LIST OF ABBREVIATIONS AND ACRONYMS		13

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013

This page intentionally left blank.

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013

1.0 SUMMARY

Intelligence testing has multiple uses in the selection and assessment of United States Air Force (USAF) pilots and pilot training candidates. Such a test of ability is a critical part of USAF medical flight screening and aeromedical waiver procedures. The purpose of this study is to assess the factorial structure regarding a computer-based intelligence test given to USAF pilot training candidates (manned as well as unmanned) during medical flight screening. Principal components analysis was conducted on the intelligence test scores from the computer-based Multidimensional Aptitude Battery-Second Edition (MAB-II). This test was administered to 10,612 USAF pilot training candidates. Subtest and measurement model correlations were estimated. Confirmatory factor analysis using this model of the MAB-II showed positive correlations between the factors and between specific subtests. The results of factor analyses suggest MAB-II intelligence test scores are best suited to a three-factor model unique to the rated USAF pilot population. In addition to verbal- and performance-based intelligence quotients, a visual reasoning and working memory intelligence quotient composed of the arithmetic, digit symbol, and spatial analyses subtests appears in test scores for this population. This finding is in contrast to the two-factor measurement model used for the general population. The results of the study support an alternative way for conceptualizing and reorganizing MAB-II intelligence test subtest scores to better account for the relationship between test scores from USAF pilot training candidates. The relationship of these test scores must be well understood to effectively evaluate how specific cognitive aptitudes are related to (and/or affected by) changes in any particular subtest. Rather than just measuring global cognitive ability, it is recommended that neuropsychologists and clinical psychologists remain sensitive to the pattern and factor analyses of subtests that comprise the global score of intelligence. This is important for fully understanding the relationship between specific cognitive aptitudes when interpreting and rendering discretionary judgments during aeromedical evaluation procedures. This is especially important when the purpose of a psychological evaluation is to assess the cognitive strengths and weaknesses of a specific individual (and his or her readiness to participate) in the unique high-risk, high-demand nature of USAF military pilot duties.

2.0 INTRODUCTION

U.S. Air Force (USAF) pilots are in a high-risk profession where mistakes in the performance of their flying duties can lead to significant costs in terms of military assets, readiness, and human life. Each year, there are several aviation-related mishaps in training and operational missions where human error is considered a causal factor. As a result of the high-risk nature of military flying and the evidence that a pilot's psychological disposition is crucial for safe and effective flying, the cognitive assessment of USAF pilots is a critical part of medical flight screening and aeromedical waiver procedures.

The assessment of cognitive functioning of USAF pilot training candidates and incumbents at the Aerospace Medicine Consultation Division within the USAF School of Aerospace Medicine includes a computer-based intelligence test, the Multidimensional Aptitude Battery – Second Edition (MAB-II) [1]. This test is integral to the (a) comprehensive medical flight screening process for training candidates prior to training, as well as (b) the comprehensive clinical and neuropsychological evaluations for rated pilots seeking an aeromedical waiver for a condition that affects their psychological disposition. The results of these tests are especially

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013

important to pilots and pilot candidates seeking a waiver for flying due to a history of an illness (i.e., bacterial meningitis, obstructive sleep apnea) or injury (i.e., traumatic brain injury) that may have affected their cognitive disposition. Performance on the tests can be compared with baseline scores taken during medical flight screening and collected prior to pilot training to assess for changes that have occurred over time and may be due to a history or presence of an aeromedically disqualifying condition. Pattern analyses of the test scores of specific subtests and indices, as well as differences between current and baseline test scores, may reflect general or specific impairments in cognitive functioning. Such analyses may also provide evidence for recovery. Regardless, intelligence testing helps flight medicine physicians assess for the cognitive suitability of a pilot applicant for training (and a rated pilot's readiness to return to flying).

However, it is critical that psychologists conducting clinical and neuropsychological aeromedical evaluations have a clear understanding of the relationship between the indices and factorial structure of the intelligence test used to make important discretionary judgments about a pilot applicant's or rated pilot's cognitive ability and readiness. Such information is helpful when interpreting and analyzing the pattern of subtest and index scores for relative strengths and weaknesses of a training candidate and making conclusions regarding how his or her profile addresses the demand of a high-risk, high-demand training and operational environment.

2.1 Conceptualizing and Measuring Intelligence

Intelligence is a global label that refers to a generalized aspect of capability and functioning assumed to produce specific responses in thought and behavior in reaction to various conditions. Intelligence is inferred from observable responses and behavioral accomplishments. As such, intelligence may be considered similar to the term "force" in physics: force is known by its effects, and its presence must be inferred from observable reactions that may be reliably tested over time.

Over the course of the last century, the understanding of observable reactions and factorial structure of abilities that comprise general intelligence continues to evolve as attention to cognitive processes and mechanisms, a deeper understanding of related issues, and new theories have emerged. Theories have evolved from conceptualizing intelligence as a single underlying factor [2,3] to multiple factor theories [4,5], to hierarchical models that describe specific abilities arranged according to increasing specificity and developmental complexity [6-8]. Theories have also evolved from strictly biological approaches [9-12] to highly complex information processing approaches [13]. Despite the notion of several competing theories, most definitions of intelligence imply, include, or elaborate on the following five areas: abstract thinking, learning from experience, solving problems through insight, adapting to new situations and information, and focusing and sustaining abilities to achieve a desired goal [14].

The most prominent theory of intelligence over the past two decades emerged from Wechsler [15]. He considered intelligence to be a global construct that was an aggregate of specific abilities involving the ability to act purposefully, think rationally, and deal effectively with the environment. He explained that intelligence is a global construct because it characterizes individual behavior as a whole. It is also specific because it is composed of elements that are qualitatively different. He also proposed intelligence should be measured by both verbal- and visual-performance-based tasks, which measure ability in different ways and are aggregated to form the general, global construct. Although factor analytic studies account for a significant

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013

percentage of the overall variance of intelligence, he also believed in another group of attributes (such as basic motivations and personality traits such as persistence) not tapped directly by existing measures of intellectual ability.

Wechsler's theory of intelligence is central to the development of the mostly widely used intelligence test in the United States, the Wechsler Adult Intelligence Scale-IV (WAIS-IV) [16]. The WAIS-IV generates a full-scale intelligence quotient (FSIQ) as a measure of the global construct of intelligence. The FSIQ is an aggregate score from several subtests. The WAIS intelligence test is the most common test of intelligence used by practitioners [17]. Earlier versions of the WAIS organized intelligence subtests into specific verbal- and visual-performance-based categories. Subtests were organized into a two-factor hierarchical model. However, changes were made to the WAIS-IV based upon factor analytic studies and revisions to subtests as well as current conceptualization of intellectual measurement based upon the theories of Carroll, Cattell, and Horn [10,18-21]. The two-factor model (verbal and visual-performance) is no longer utilized and has been replaced with a four-factor model (i.e., verbal comprehension, perceptual reasoning, working memory, and processing speed). Each index contributes to the FSIQ. However, the changes in the WAIS-IV raise questions as to whether factor analytic studies support changes to the factor analytic structure of earlier versions used with specific populations.

As mentioned previously, the MAB-II is a computerized intelligence test. This test is based upon the theory, content, and factor analytic studies of earlier versions of the WAIS. A previous study evaluating MAB-II scores among USAF pilot training candidates identified the relationships between subtests and provided confirmatory analyses for a two-factor hierarchical model predicated on earlier versions of the WAIS [22]. Additionally, normative data for USAF pilot training candidates and their MAB-II test scores have been well documented [23].

However, it is unknown if the MAB-II factorial structure of the subtest test scores for USAF pilot training candidates is organized in the same fashion as the general population. To date, factor analytic studies assessing for alternative hierarchical models have not been conducted on the MAB-II in regards to pilot applicant test scores. The MAB-II factor structure was designed on the earlier versions of the WAIS, and based upon modernized changes to the WAIS and conceptualization of intelligence, it is reasonable to evaluate the efficacy of a two-factor hierarchical model. Identification of a three-, four-, and five-factor model may help improve theoretical understanding and organization of subtests accounting for the variance of global index intelligence test scores.

Because of the importance of intelligence testing in aeromedical clinical and neuropsychological evaluations of pilots, the sensitive nature of the aeromedical waiver process, and the implications for mission readiness and safety, it is essential to have an accurate and clear understanding of the statistical relationships between subtest scores on this test and the factors that account for the variance within the global FSIQ construct for such a unique occupational group.

2.2 Purpose of the Study

The purpose of this study is to investigate the goodness-of-fit for two-, three-, four-, and five-factor measurement models of the MAB-II for USAF pilot training applicant population data. The results of the study aim to improve the clinical acumen of psychologists who must

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013

analyze the pattern of subtest scores when evaluating USAF pilot training applicants and rated incumbents.

3.0 METHODS

3.1 Participants

In total, 10,612 pilot candidates who went through medical flight screening over the past 10 years were included in the study. Participants, at the time of testing, had a mean age of 22 years (standard deviation (SD) = 2.7); 84% were Caucasian, 4% were Hispanic, 2% were African-American, 6% were “other,” and 4% did not report their ethnicity. Ninety-one percent were male and 9% were female. All were either college graduates or enrolled in their fourth year of college at the time of testing. All participants were found to be physically and psychologically healthy and to have met the enhanced aeromedical standards required for attending pilot training and becoming a rated USAF pilot.

3.2 Instrument

The MAB-II is a broad-based test of cognitive functioning with well-documented internal consistency, test-retest reliability, and validity coefficients [1]. The content and structure of the test were fashioned after the WAIS-III [24], which is a widely used (and most preferred), individually administered test of cognitive functioning and intelligence among neuropsychologists [17]. The MAB-II has 10 subtests that are each 7 minutes long, and all items have five multiple-choice responses. The MAB-II requires only 70 minutes to complete and can be administered in group settings. Administration of this test produces verbal intelligence quotient (VIQ), performance intelligence quotient (PIQ), and FSIQ scores, which are measures of cognitive functioning. The test is separated into verbal abilities (i.e., subtests of information, comprehension, arithmetic, similarities, and vocabulary) and performance abilities (i.e., digit symbol coding, picture completion, spatial analyses, picture arrangement, and object assembly). MAB-II normative subtest scores for the general population have a mean of 50 and an SD of 10. The VIQ, PIQ, and FSIQ scores in the general population have a mean of 100 and an SD of 15. The mean USAF pilot training applicant MAB-II VIQ, PIQ, and FSIQ scores range between 119 and 121, with SDs between 6 and 8, respectively [22,23,25]. The MAB-II manual has well-documented internal consistency, test-retest reliability, and validity coefficients. For a description of the subtest (abbreviations) and factors measured, see Table 1 for more information.

3.3 Procedure

The sample of USAF pilot training candidates in this study was administered the MAB-II as a routine part of medical flight screening prior to attending pilot training. The test was administered in an access-controlled classroom under a standardized protocol based upon manual instructions. Training candidates were briefed on the purpose and potential uses of their test scores for aeromedical waiver requirements. Environmental aspects of the classroom (i.e., temperature, noise, interruptions) were monitored and controlled to provide an optimal testing environment.

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013

Table 1. MAB-II Index and Subtest Descriptions

Subtest	Description
	VIQ
Information (inf)	General fund of knowledge; long-term memory
Comprehension (com)	General social reasoning and comprehension
Arithmetic (ari)	Numerical reasoning and problem solving
Similarities (sim)	General reasoning and problem solving
Vocabulary (voc)	Flexibility and adjustment to novelty, reasoning, abstract thought, long-term memory
	PIQ
Digit Symbol (ds)	Adaptation to new set of demands; visual learning and coding, figural memory, and speed of information processing
Picture Completion (pc)	Visual attention to detail; knowledge of common objects; perceptual and analytical skills
Spatial Analyses (sp)	Ability to visualize and mentally rotate abstract two-dimensional images of objects in different positions; figural-domain reasoning
Picture Arrangement (pa)	Visual reasoning; ability to identify a meaningful sequence; social intelligence; perceptual reasoning
Object Assembly (op)	Visualization and visuo-construction skills; perceptual analytical skills needed to identify a meaningful object from left-to-right sequence

The test data were scored and downloaded into a secure server maintained by information technology staff. The data were loaded into individual electronic files for each training candidate and checked for errors, omissions, and other problematic issues as part of data quality control procedures. The data were examined for missing, out of range, or inappropriate values. Participants with erroneous data were removed. Twelve subjects were removed for this reason. Principal components analyses and confirmatory factor analyses were conducted using AMOS 17.0 [26,27]. Estimation was accomplished using covariance matrices and maximum likelihood methods.

Two-, three-, four-, and five-factor structures of the MAB-II were calculated for the pilot applicant population. Goodness-of-fit indices for the models were defined as having a goodness-of-fit > 0.90, comparative fit index > 0.90, and root mean square error of approximation < 0.08 inclusive [28]. Those models that did not meet these criteria were dismissed. Remaining models with sufficient goodness-of-fit were compared to each other. To test for significant difference between two models, χ^2 was estimated for each. The difference between the χ^2 values and the difference in the degrees of freedom for the two models was evaluated at $p < 0.05$ [27]. A significant χ^2 would indicate a difference in the two models and require the more complex model be used. A nonsignificant χ^2 would allow the less complex model to be used. The correlations of all the factors were then estimated.

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013

4.0 RESULTS

The model fit suggested by the publisher of the MAB-II [1] consists of two correlated factors interpreted as verbal (VIQ) and performance (PIQ). The verbal consists of the following subtests: information (inf), comprehension (com), arithmetic (ari), similarities (sim), and vocabulary (voc). Performance consists of digit symbol (ds), picture completion (pc), spatial analyses (sp), picture arrangement (pa), and object assembly (op). Principal components analysis constrained to a two-factor model for the pilot population revealed the same indices (Table 2). Goodness-of-fit criteria were met, and confirmatory factor analysis of this model is shown in Figure 1.

Table 2. Two- and Three-Factor Models of the MAB-II

Two-Factor Model	Three-Factor Model
VIQ	
Information	Information
Comprehension	Comprehension
Arithmetic	
Similarities	Similarities
Vocabulary	Vocabulary
PIQ	
Digit Symbol	
Picture Completion	Picture Completion
Spatial Analyses	
Picture Arrangement	Picture Arrangement
Object Assembly	Object Assembly
VRMIQ	
	Arithmetic
	Digit Symbol
	Spatial Analyses

Similarly, a three-factor model also met goodness-of-fit criteria. The third factor in this model, visual reasoning and working memory IQ (VRMIQ), consists of arithmetic from the VIQ and digit symbol and spatial analyses from the PIQ in the two-factor model as shown in Table 1. Confirmatory factor analysis of this model is shown in Figure 2. Four- and five-factor models did not meet the goodness-of-fit criteria and were discounted. Goodness-of-fit statistics for the two- and three-factor models are shown in Table 3. To determine whether the three-factor model was superior to the two-factor model, χ^2 analyses were conducted. The χ^2 for the two-factor model was 459.826 with 26 degrees of freedom, and χ^2 for the three-factor model was 364.108 with 23 degrees of freedom. The difference of 95.718 with 3 degrees of freedom was tested at $p < 0.05$ and found to be statistically significant ($p < 0.0001$). Therefore, the more complex model with the superior fit—the three-factor model—is preferable.

Standards suggested by Cohen [29] divide correlations into three groups based on Cohen's *d*, a measure of effect size. Correlations categorized as small ($r = 0.10$ to 0.23) have effect sizes of 0.20 to 0.49, moderate correlations ($r = 0.24$ to 0.36) have effect sizes of 0.50 to 0.79, and large correlations ($r >= 0.37$) have effect sizes of 0.80 or greater. Large correlations were found between the subtests and the indices to which they belong for both models evaluated. Large correlations were also found between the indices themselves with the exception of the moderate correlation between the VIQ and VRMIQ in the three-factor model. This suggests a commonality of measured constructs.

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013

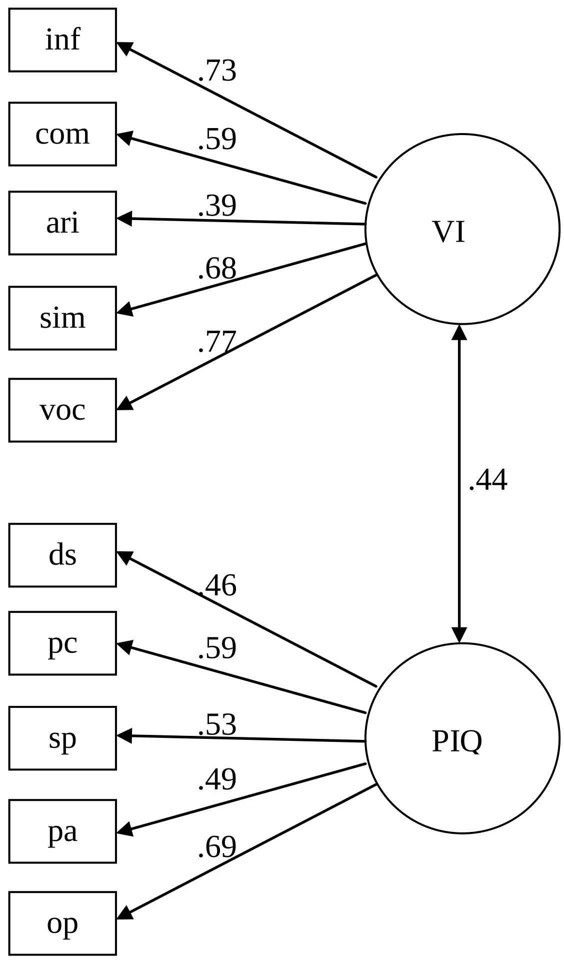

Figure 1. Confirmatory Factor Analysis of the MAB-II Two-Factor Model

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013

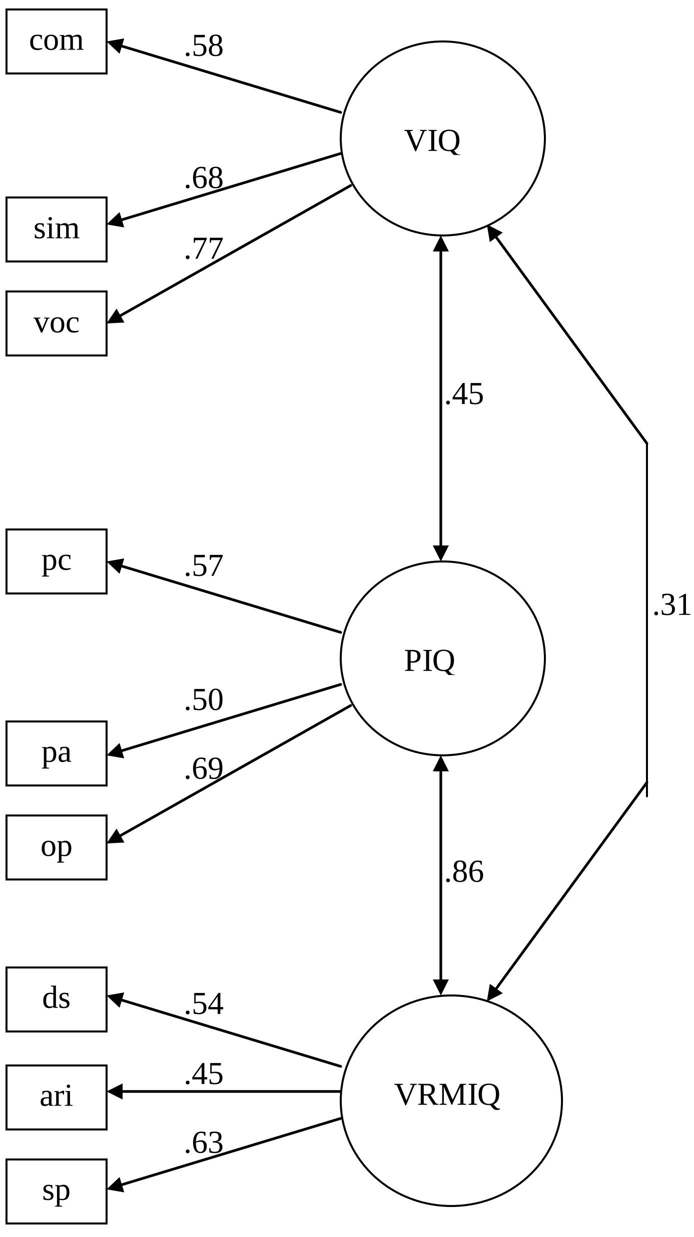

Figure 2. Confirmatory Factor Analysis of the MAB-II Three-Factor Model

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013

Table 3. Goodness-of-Fit Statistics for the Two Models

Model	χ^2	Degrees of Freedom	GFI[a]	CFI[a]	RMSEA[a]
Two Factor	459.826	26	0.991	0.981	0.040
Three Factor	364.108	23	0.993	0.985	0.037
Difference	95.718	3			

[a]GFI = goodness-of-fit index; CFI = comparative fit index; RMSEA = root mean square error of approximation

5.0 DISCUSSION

The results of factor analyses reveal a three-factor hierarchical model is a more effective and suitable model for organizing subtests accounting for the variance of global intelligence index test scores from USAF pilot training candidates. Instead of the two-factor model consisting of a VIQ and PIQ (as applied to the general population and earlier versions of the WAIS), test scores from the USAF pilot training candidates appear more suited for a three-factor model that adds a VRMIQ index.

The VRMIQ consists of the subtests of arithmetic, digit symbol, and spatial analyses subtests, indicating that each subtest shares the components of reasoning and memory. The arithmetic subtest is defined as general numerical reasoning and problem solving. Although this subtest is correlated with VIQ based upon general population normative data, it is included as a component of the VIQ factor.

When evaluating arithmetic subtest scores for USAF pilots, this ability appears to be more associated with working memory and reasoning processes measured by the digit symbol and spatial analyses subtest scores than the subtests within the VIQ. Similarly, USAF pilot applicant test scores on digit symbol (measuring visual learning and coding and speed of information processing) and spatial analyses (measuring the ability for visual reasoning via mental rotation of abstract two-dimensional images of objects) appear to be more suited under a factor separate from the general performance-based IQ index.

The MAB-II is a critical component to medical flight screening of USAF pilot candidates as well as aeromedical evaluation processes for USAF rated pilots. The results of this study serve to improve the clinical acumen of neuropsychologists' understanding of the relationship between MAB-II subtest scores and indices for this unique group. In particular, the results of this study provide a level of scrutiny that allow a neuropsychologist to effectively understand pattern analyses of subtest scores by determining which aptitudes share a significant level of interrelatedness. For example, the findings from the three-factor analytic study indicate the abilities of numerical and spatial reasoning and working memory represent a singular, complex factor independent of traditional VIQ and PIQ factors. Thus, pilot training candidates or rated pilot incumbents who have deficits in numerical reasoning and working memory may be at increased likelihood to experience difficulties with spatial reasoning related to the mental rotation of two-dimensional images of objects to different positions. These correlations are meaningful for clinical interpretation of test scores and reflect the relationship between cognitive abilities logically relevant to flying.

The changes in the factor analytic organization of subtests for the WAIS-IV raise questions as to whether or not research would also change the factor analytic structure of subtest scores for the MAB-II. The WAIS-IV organizes arithmetic, spatial analyses, and digit symbol subtests into the indices of working memory, perceptual reasoning, and processing speed, respectively. The results of this study suggest the three-factor analytic model of the MAB-II is a

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013

suitable method for organizing the subtests of arithmetic, digit symbol, and spatial analyses subtests. The results appear more similar with the way the WAIS-IV organizes such subtests. Because all subtests on the MAB-II are timed, processing speed is an overarching aptitude measured throughout the entirety of the testing and a component built in to current indices.

Baseline intelligence testing with the MAB-II is used, under various circumstances, to identify pilot training candidates who may be at high risk of performance difficulties. The findings from the three-factor model (adding a visual reasoning and working memory index) reveal additional insight into the cognitive abilities and profiles of USAF pilot training candidates. Furthermore, the results of the study provide guidance into the direction of future studies targeting cognitive functioning and pilot performance outcomes.

However, when evaluating the outcomes and making comparisons between the MAB-II and WAIS-IV regarding individual and/or group scores, there are several differences to consider. Although an exhaustive list of such differences is beyond the scope of this study, significant considerations include the following:

(a) Discrepancies and variation between computer-based intelligence testing of the MAB-II and the supposed corresponding subtests on the WAIS-IV
(b) Changes in the delivery of test items may result in measurement of slightly different abilities
(c) Theoretical organization of variables based upon developments in the way intelligence is conceptualized
(d) MAB-II multiple choice item response format versus WAIS-IV open response format
(e) Visual performance subtests having a psychomotor and dexterity component
(f) All MAB-II item responses are timed versus a limited and specific of set of items being timed on the WAIS –IV
(g) All or nothing credit for responses on the MAB-II versus items on the WAIS-IV that allow for partial credit

Such differences should be taken into account when making any sort of comparisons and attempts to generalize scores between tests.

6.0 CONCLUSIONS

There are significant differences between the general population and rated USAF pilot intelligence test scores. The MAB-II is regularly used as a part of the evaluation process for selecting applicants for pilot training and for considering a pilot’s readiness for returning to flying duties after being disqualified following a neurological or psychological insult. This study demonstrates that a three-factor model for the MAB-II is better suited for the USAF pilot population than the two-factor model based upon the general population. The relationship of these test scores must be well understood for effective neuropsychological and clinical psychological specialty evaluations of USAF pilots.

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013

7.0 REFERENCES

1. Jackson DN. Multidimensional Aptitude Battery-II: manual. Port Huron, MI: Sigma Assessment Systems, Inc.; 2003.

2. Spearman C. "General intelligence," objectively determined and measured. Am J Psychol 1904; 15, 201-9.

3. Spearman C. The abilities of man: their nature and measurement. New York: Macmlllan; 1927.

4. Thurstone LL. Primary mental abilities. Chicago: University of Chicago Press; 1938.

5. Gardner H. Multiple intelligences: the theory in practice. New York: Basic Books; 1993.

6. Guilford JP. Some changes in the structure-of-intellect model. Educational and Psychological Measurement 1988; 48(1):1-4.

7. Horn JL. Remodeling old models of intelligence. In: Wolman BB, ed. Handbook of intelligence: theories, measurements, and applications. Hoboken, NJ: Wiley; 1985:267-300.

8. Vernon PE. The structure of human abilities, 2nd ed. London: Methuen; 1961.

9. Halstead WE. Biological intelligence. In: Jenkins JJ, Paterson DG, eds. Studies in individual differences: the search for intelligence. New York: Appleton-Century-Crofts; 1961:661-8.

10. Horn JL, Cattell RB. Refinement and test of the theory of fluid and crystallized intelligences. J Educ Psychol 1966; 57(5):253-70.

11. Hebb DO. Textbook of psychology, 3rd ed. Philadelphia: W.B. Saunders; 1972.

12. Piaget J. The psychology of intelligence. New York: Harcourt, Brace & World; 1950.

13. Sternberg RJ. Beyond IQ: a triarchic theory of human intelligence. New York: Cambridge University Press; 1985.

14. Groth-Marnat G. Handbook of psychological assessment, 3rd ed. New York: John Wiley & Sons, Inc.; 1997:133.

15. Wechsler D. The measurement and appraisal of adult intelligence. Baltimore, MD: Williams & Wilkins; 1958.

16. Wechsler D. Wechsler Adult Intelligence Scale-fourth edition. Technical and interpretive manu

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013

17. Rabin LA, Barr WB, Burton LA. Assessment practices of clinical neuropsychologists in the United States and Canada: a survey of INS, NAN, and APA Division 40 members. Arch Clin Neuropsychol 2005; 20(1):33-65.

18. Carroll JB. Human cognitive abilities: a survey of factor-analytic studies. Cambridge, England: Cambridge University Press; 1993.

19. Carroll JB. The higher-stratum structure of cognitive abilities: current evidence supports g and about ten broad factors. In: Nyborg H, ed. The scientific study of general intelligence: tribute to Arthur R. Jensen. Oxford, England: Elsevier Science Ltd.; 2003:5-21.

20. Cattell RB, Horn JL. A check on the theory of fluid and crystallized intelligence with description of new subtest designs. Journal of Educational Measurement 1978; 15:139-164.

21. Horn JL. Measurement of intellectual capabilities: A review of theory. In: McGrew K, Werder JK, Woodcock RW. WJ-R technical manual. Rolling Meadows: Riverside Publishing; 1991:197-232.

22. Chappelle W, Ree MJ, Barto EL, Teachout MS, Thompson WT. Joint use of the MAB-II and MicroCog for improvements in the clinical and neuropsychological screening and aeromedical waiver process of rated USAF pilots. Brooks City-Base, TX: U.S. Air Force School of Aerospace Medicine; 2010 Jan. AFRL-SA-BR-TR-2010-0002.

23. King RE, Barto E, Ree MJ, Teachout MS, Retzlaff P. Compilation of pilot cognitive ability norms. Wright-Patterson AFB, OH: U.S. School of Aerospace Medicine; 2011 Dec. AFRL-SA-WP-TR-2012-0001.

24. Wechsler D. Wechsler Adult Intelligence Scale-3rd edition (WAIS-3®). San Antonio, TX: Harcourt Assessment; 1997.

25. King RE, Retzlaff P, Barto E, Ree MJ, Teachout MS, Carretta TR. Pilot cognitive functioning and training outcomes. Wright-Patterson AFB, OH: U.S. Air Force School of Aerospace Medicine; 2012 Jul. AFRL-SA-WP-TR-2012-0004.

26. Arbuckle JL. Amos™ 17.0 user's guide. Chicago: SPSS Inc.; 2008.

27. Blunch N. Introduction to structural equation modeling using SPSS and Amos. Thousand Oaks, CA: Sage Publications Ltd.; 2008.

28. Lance CE, Vandenberg RJ. Statistical and methodological myths and urban legends: doctrine, verity and fable in the organizational and social sciences. New York: Routledge; 2009.

29. Cohen J. Statistical power analysis for the behavioral sciences, 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.; 1988.

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013

LIST OF ABBREVIATIONS AND ACRONYMS

FSIQ	full-scale intelligence quotient
MAB-II	Multidimensional Aptitude Battery-Second Edition
PIQ	performance intelligence quotient
SD	standard deviation
USAF	U.S. Air Force
VIQ	visual intelligence quotient
VRMIQ	visual reasoning and working memory intelligence quotient
WAIS	Wechsler Adult Intelligence Scale

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2013-1912, 22 Apr 2013